

# Optimal Placement Algorithms for Virtual Machines

Umesh Bellur

Department of Computer Science  
Indian Institute of Technology Bombay  
Mumbai 400076, India

Chetan S. Rao

Department of Computer Science & Engineering  
National Institute of Technology Calicut  
Kerala 673601, India

Madhu Kumar S.D.

Department of Computer Science & Engineering  
National Institute of Technology Calicut  
Kerala 673601, India

November 24, 2010

## Abstract

Cloud computing provides a computing platform for the users to meet their demands in an efficient, cost-effective way. Virtualization technologies are used in the clouds to aid the efficient usage of hardware. Virtual machines (VMs) are utilized to satisfy the user needs and are placed on physical machines (PMs) of the cloud for effective usage of hardware resources and electricity in the cloud. Optimizing the number of PMs used helps in cutting down the power consumption by a substantial amount.

In this paper, we present an optimal technique to map virtual machines to physical machines (nodes) such that the number of required nodes is minimized. We provide two approaches based on linear programming and quadratic programming techniques that significantly improve over the existing theoretical bounds and efficiently solve the problem of virtual machine (VM) placement in data centers.

# 1 Introduction

Cloud computing is a large scale network-based distributed computing environment where computing resources such as memory, processing power, bandwidth, etc. are available on demand to the users. The cloud computing environment comprises of many models such as *Software as a Service (SaaS)*, *Platform as a Service (PaaS)* and *Infrastructure as a Service (IaaS)*. These models are made available to the users through virtualization techniques. The users' demands are satisfied by a set of servers hosted on virtual machines (VMs). The VMs utilize the resources of underlying physical machines (PMs) or nodes provided and operated by organizations called 'cloud providers'. Some examples of cloud providers include Amazon EC2 [1], GoGrid [2] and Rackspace Cloud [3].

Adopting the use of virtual machines (VMs) in such large-scale environments enhances the number of available servers through multiple OS instances on a single node, thereby achieving efficient hardware utilization. However, there may be a number of underutilized nodes due to the inefficient mapping of virtual machines to physical machines. Minimizing the number of physical machines utilized helps in cutting down the power consumption drastically [9].

The placement algorithm for VMs in a data center allocates various resources such as memory, bandwidth, processing power, etc. from a physical machine (PM) to VMs such that the number of PMs used is minimized.

This problem can be viewed as a multi-dimensional packing problem [15] (Figure 1). The resource requests of VMs are considered as  $d$ -dimensional vectors with non-negative entries (balls). The resource available at each PM is considered to be a  $d$ -dimensional vector (each dimension signifies an independent resource) with a magnitude of 1 along each dimension (bins). The goal is to minimize the number of bins such that for every bin the sum of the vectors placed in that bin is coordinate-wise no greater than the bin's vector. Thus, the resource allocation problem is an instance of the  $d$ -dimensional Vector Bin Packing problem (VBP) [15].

For  $d = 1$ , the VBP is identical to the 1-dimensional Bin Packing problem.

We now define the optimization problem that we are addressing in this paper.

## **Vector Bin Packing problem (VBP)**

*Given a set  $S$  of ' $n$ '  $d$ -dimensional vectors  $p_1, p_2, \dots, p_n$  from  $[0, 1]^d$ , find a packing (partition) of  $S$  into  $A_1, A_2, \dots, A_m$  such that  $\sum_{p \in A_i} p^k \leq 1, \forall i, k$*

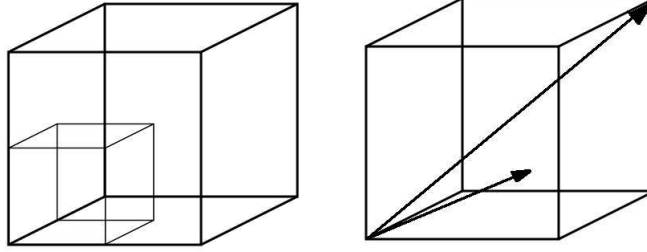


Figure 1: Bin-packing & Vector Bin Packing along 3-dimensions

( $p^k$  denotes the projection of vector  $p$  along ' $k$ 'th dimension). The objective is to minimize the value of ' $m$ ', the number of partitions.

The vector bin packing problem is a computationally hard problem and it is known to be NP-Hard [17].

## 2 Related work

**Vector Bin Packing (VBP)** One dimensional bin packing problem has been studied extensively. Fernandez de la Vega and Lueker [14] gave the first Asymptotic Polynomial-Time Approximation Scheme (APTAS). They put forward a rounding technique that allowed them to reduce the problem of packing large items to finding an optimum packing of just a constant number of items (at a cost of  $\epsilon$  times the optimal solution - OPT). Their algorithm was later improved by Karmarkar and Karp [21], to a  $(1+\log^2)$ -OPT bound.

For 2-dimensional vector bin packing, Woeginger [32] proved that there is no APTAS. For higher dimensions, Fernandez de la Vega and Lueker [14] proposed a simple  $(d + \epsilon)$ -OPT algorithm, which extends the idea of 1-dimensional bin packing. Chekuri and Khanna [12] showed an  $O(\log d)$ -approximation algorithm that runs in polynomial time for fixed  $d$ . Bansal et al. [6] improved this result, showing a  $(\ln d + 1 + \epsilon)$ -approximation algorithm for any  $\epsilon \geq 0$ . Karger et al. [20] have recently proposed a polynomial approximation scheme for randomized instances of the multidimensional vector bin packing using *smoothing techniques*. Patt-Shamir et al. [26] have recently explored the vector bin packing problem with bins of varying sizes and propose a  $(\ln 2d + 1 + \epsilon)$ -approximation algorithm for any  $\epsilon \geq 0$ .

**Placement Algorithm** The problem of VM placement is at the core of cloud computing. Several research works address the importance of placing VMs appropriately [16, 9, 7]. Vogels [31] quotes the benefit of packing VMs efficiently in server consolidation. Recently, Hermenier et al. [18] developed a constraint programming based mechanism for dynamic consolidation.

Several modified versions of First-Fit Decrease (FFD) have been used for VM placements. Verma et al. [30] propose an algorithm to pack VMs optimally while minimizing the number of migrations. Khanna et al. [24] propose a reconfiguration algorithm to cut down the wastage of physical resources. Hyser et al. [19] propose an iterative rearrangement technique for improving placements in a dynamic scenario. Bobroff et al. [5] presents a dynamic algorithm that forecasts the resource demands and packs VMs. Shahabuddin et al. [28] propose a simple heuristic which aims to efficiently allocate resources.

Despite the recent research trends towards virtualization, the problem of VM placements is vastly unexplored. To overcome this limitation, we propose a linear programming based approach which places VMs efficiently on a set of PMs.

The rest of this paper is organized as follows. Section 3 deals with the formulation of the problem, Section 4 provides our algorithms for vector bin packing (VBP) and Section 5 describes the experimental setup and results. Section 6 concludes the paper.

### 3 Problem formulation

We formulate the problem as an integer program in subsection 3.1. The integer constraints are relaxed and we formulate it's dual (in subsection 3.2). The solution of the relaxed integer linear program gives a thoughtful insight about the optimal number of bins.

#### 3.1 Integer Linear Program (ILP) formulation

The vector bin packing problem (VBP) can be formulated as an integer program. We use two binary variables  $x_{ij}$  and  $y_j$ . The binary variable  $x_{ij}$  indicates if vector  $p_i$  is assigned to bin  $j$  and the binary variable  $y_j$  indicates whether bin  $j$  is in use or not. Our objective is to minimize the number of bins used.

The number of bins  $m$  can initially be set to a sufficiently large value arrived at by any heuristic (example - de la Vega and Leuker [14] give a

Notation Table	
$x_{ij}$	Fraction of vector $i$ packed in $j$ th bin
$y_j$	Binary variable to determine usage of bin $j$
$p_i$	Input vector $i$ (VM $_i$ )
$n$	Number of vectors (VMs)
$m$	Number of bins (PMs)
$d$	Dimension of each vector

Figure 2: Notation table for the integer linear program (ILP) formulation

$O(d)$ -OPT bound on the number of bins). Then, we formulate the integer program (ILP) as follows -

$$\text{minimize : } \sum_j y_j \quad \text{s.t.} \quad (1)$$

$$\sum_j x_{ij} = 1 \quad 1 \leq i \leq n \quad (2)$$

$$\sum_i p_i^k \cdot x_{ij} \leq 1 \quad 1 \leq j \leq m, 1 \leq k \leq d \quad (3)$$

$$y_j \geq x_{ij} \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (4)$$

$$x_{ij} \in \{0, 1\} \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (5)$$

The notations are mentioned in Figure 2. The constraints of the ILP are as follows -

- Constraint (2) states that every vector is packed in a bin.
- Constraint (3) ensures that the packed vectors do not exceed the bin dimensions.
- Constraint (4) tells whether a bin is used or not.
- Constraint (5) ensures that a vector is either packed entirely in a bin or not.

Constraint (5) can be relaxed as follows to obtain a linear program (LP).

$$x_{ij} \geq 0 \quad 1 \leq i \leq n, \quad 1 \leq j \leq m \quad (5a)$$

We can obtain a feasible solution for the LP using any standard method [13]. Using binary search technique, we can also find the least value of  $m, m' \in Z^+$  for the relaxed ILP for which a feasible solution exists. The value of  $m'$  thus obtained will be less than the optimal solution for the integer program i.e. ( $m' \leq \text{OPT}$ ). However, the solution obtained is usually not integral. To tackle this problem, we formulate a dual-maximization problem [29] for the above relaxed ILP.

### 3.2 Dual-maximization problem

We introduce several new variables  $z_{ij}$  to formulate the dual. The dual-maximization problem formulation is given in the Appendix 8.1. We arrive at the following set of equations and constraints -

$$\text{maximize : } \sum_i \sum_j x_{ij} z_{ij} \quad \text{s.t.} \quad (6)$$

$$\sum_j x_{ij} = 1 \quad 1 \leq i \leq n \quad (7)$$

$$\sum_i p_i^k . x_{ij} \leq 1 \quad 1 \leq j \leq m, 1 \leq k \leq d \quad (8)$$

$$\sum_i z_{ij} \leq 1 \quad 1 \leq j \leq m \quad (9)$$

$$x_{ij}, z_{ij} \geq 0 \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (10)$$

This is a nonlinear program (NLP) as the objective function is nonlinear. Hereafter, we shall refer to it as NLP.

The number of variables in the NLP can be reduced by performing the following substitutions for the value of  $z_{ij}$ 's.

**Theorem 1.** *The optimal solution to the NLP will still be optimal when the value of  $z_{ij}$  is replaced by  $x_{ij} / \sum_i x_{ij}$ .*

PROOF. From the Jensen's Inequality, we have that if  $f$  is a convex function ("concave-up") on an interval  $I$  and  $a_i \in I$  then for weights  $\lambda_i$  summing to 1 -

$$f\left(\sum_{i=1}^n \lambda_i a_i\right) \leq \sum_{i=1}^n \lambda_i f(a_i)$$

We can apply Jensen's inequality with  $\lambda_i$  and  $a_i$  corresponding to  $x_{ij}$  and  $z_{ij}$ , respectively. The modified set of equations in this case is as follows-

$$f\left(\sum_{j=1}^m x_{ij} z_{ij}\right) \leq \sum_{j=1}^m x_{ij} f(z_{ij}) \quad 1 \leq i \leq n \quad (\text{a})$$

From the property of convex functions, we have -

$$f(tx) \leq tf(x) \quad 0 \leq t \leq 1 \quad (\text{b})$$

From (a) and (b), we have -

$$f\left(\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m x_{ij} z_{ij}\right) \leq \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m x_{ij} f(z_{ij}) \quad (\text{c})$$

Since  $f(x)$  is a convex function, any value which maximizes  $x$  also maximizes  $f(x)$  and vice-versa. Hence, from the inequality (c), we have that the term  $x_{ij}f(z_{ij})$  should be maximized for the objective function to be maximized. Indirectly,  $z_{ij}$  has to be maximized relative to the values of  $x_{ij}$ . The value of  $z_{ij}$  is constrained by the constraint (9), and hence we come up with the following tight function for  $z_{ij}$  -

$$\begin{aligned} z_{ij} &= \frac{x_{ij}}{\sum_i x_{ij}} \\ \implies \sum_i z_{ij} &= \sum_i \frac{x_{ij}}{\sum_i x_{ij}} = 1 \end{aligned} \quad (11)$$

■

Since  $\sum_i \sum_j x_{ij} = n$ , the value of  $\sum_i x_{ij} \approx n/m$  for an appropriately chosen value  $m$ . Thus, the objective function can be reduced to a quadratic term.

From (11), the NLP now becomes -

$$\text{maximize : } \sum_i \sum_j x_{ij}^2 \quad \text{s.t.} \quad (12)$$

$$\sum_j x_{ij} = 1 \quad 1 \leq i \leq n \quad (13)$$

$$\sum_i p_i^k \cdot x_{ij} \leq 1 \quad 1 \leq j \leq m, 1 \leq k \leq d \quad (14)$$

$$x_{ij} \geq 0 \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (15)$$

The optimal solution to the above modified NLP - NLP' - must be necessarily integral (follows from Theorem 1). In this light, we now present our algorithms which will provide the (near-)optimal integer solution.

## 4 Algorithms and their complexity

In this section, we provide two algorithms to solve the vector bin packing (VBP) problem. The main idea is to harness the polynomial-time solvability of linear and quadratic programming techniques.

### 4.1 Quadratic programming

The quadratic program  $NLP'$  can be solved using various efficient techniques such as interior points, active set [25], gradient techniques or through the extensions of simplex algorithm [25].

**Complexity** Kozlov et al. [22] presented a polynomial time algorithm for solving convex quadratic programs. Since our objective function is a convex function,  $NLP'$  can be solved in polynomial time.

### 4.2 Linear programming

The relaxed version of the integer linear program (ILP) can also be used to derive (near-)optimal solutions for the VBP problem. The algorithm is as follows -

---

**Algorithm 1** PACKINGVECTORS( $P_n, d$ )

---

**Require:** A set of vectors  $p_1, p_2, \dots, p_n$ ;  $P_n$ .

Dimension of vectors  $d$

```
1:  $(m, X) = \text{SolveLP}(P_n, d)$ 
2: if  $m \geq \frac{n}{2}$  then
3:   return  $\text{FirstFit}(P_n, d)$ 
4: else if  $m \leq \sqrt{\frac{n}{d}}$  then
5:   return  $\text{GreedyLP}(P_n, X, d)$ 
6: else
7:   return  $\text{IterativePack}(P_n, X, d)$ 
8: end if
```

---

Algorithm 1 is an iterative algorithm which packs vectors in every iteration until the input is exhausted. The algorithm branches into 3 cases depending upon the solution returned by the relaxed integer linear program - LP.



---

**Algorithm 2** GREEDYLP( $P_n, X, d$ )

---

**Require:** A set of vectors  $p_1, p_2, \dots, p_n$ ;  $P_n$  and a set of  $x_{ij}$  values  $X$ .

```
1:  $X' = \text{SortDescending}(X)$ 
2: while  $X' \neq \Phi$  do
3:   Remove the top element  $x_{ij}$  in  $X'$ 
4:   if vector  $p_i$  fits in bin  $j$  then
5:      $\text{Pack}(i, j)$ 
6:     Remove  $p_i$  from  $P_n$ 
7:   end if
8: end while
9:  $\text{PackingVectors}(P_n, d)$ 
```

---

Algorithm 2 is a subroutine of Algorithm 1 which packs the vectors greedily, given the solution set  $X = \{x_{ij} \forall i, j\}$

---

**Algorithm 3** ITERATIVEPACK( $P_n, X, d$ )

---

**Require:** A set of vectors  $p_1, p_2, \dots, p_n$ ;  $P_n$  and a set of  $x_{ij}$  values  $X$ .

```
1:  $P'_n = P_n$ 
2:  $Z = \text{FindDualObj}(X, d)$ 
3: for  $j = 1$  to  $m$  do
4:   if  $\sum_i x_{ij} z_{ij} \geq \frac{1}{2}$  then
5:      $X'_j = \text{SortDescending}(X_j)$ 
6:      $X''_j = \text{RemoveLessThanHalf}(X'_j)$ 
7:      $\text{Pack}(X''_j)$ 
8:      $P'_n = P'_n \setminus \text{PackedVectors}$ 
9:   end if
10: end for
11:  $\text{PackingVectors}(P'_n, d)$ 
```

---

Algorithm 3 is a subroutine of Algorithm 1 which packs the bins having utility factor ( $\sum_i x_{ij} z_{ij}$ ) more than half. It ensures that the efficiently assigned vectors are packed into their corresponding bins.

**Complexity** The subroutines of Algorithm 1 - Algorithms 2,3 - run in polynomial time. Solving the relaxed integer linear program can be done in polynomial time [27]. The First Fit heuristic also runs in polynomial time. Thus, Algorithm 1 runs in polynomial time.

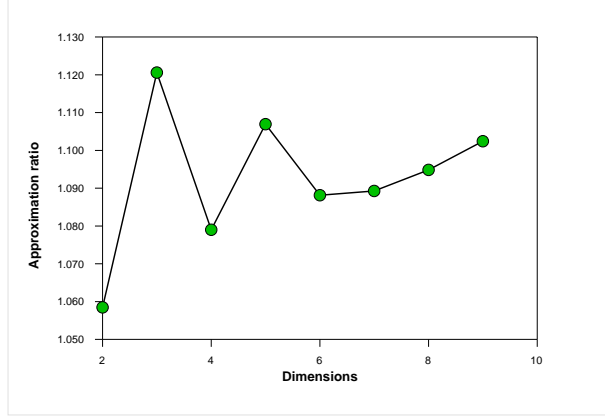


Figure 3: Mean approximation ratio - PACKINGVECTORS/OPT. Mean ratio of about 2000 randomized trials along each dimension. For dimensions  $d < 10$ , the mean approximation ratio stays below 1.2

## 5 Experimental setup and results

We test our ‘PACKINGVECTORS’ algorithm (Algorithm 1) discussed above with the existing theoretical worst-case bound for the vector bin packing (VBP) problem.

**Tools used** The Mixed Integer Linear Programming (MILP) solver ‘`lp-solve`’ [4] was used to derive exact solutions for randomized input instances (20 VM configurations). ‘`lp-solve`’ was also used as a linear program solver in Algorithm 1.

2000 iterations of randomized test inputs were performed for each dimension ranging from 2 to 10 ( $2 \leq d \leq 10$ ). The number of input VMs,  $n$ , were about 20 in each iteration.

Our results were compared with the exact solution of the optimal number of PMs (bins), and the mean approximation factor was computed. The mean approximation ratios are as shown in Figure 3.

Our results were also compared with the existing bounds of approximation given by Bansal et al. [6] and was found to have a substantial improvement as seen in Figure 4.

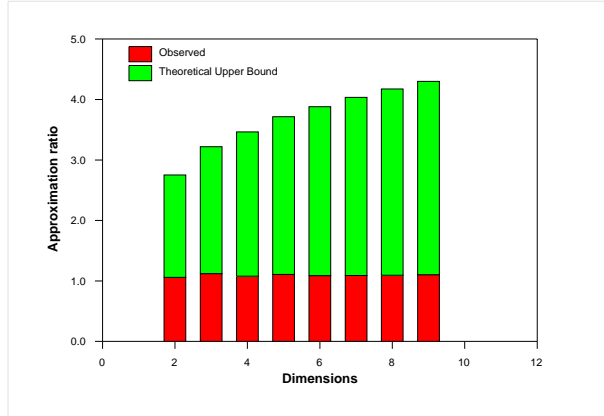


Figure 4: Our result vs. theoretical upper bound ( $\ln d$ ). The bars colored red indicate the mean approximation ratios of our algorithm whereas the green bars indicate the performance of the current best algorithm [6]

## 6 Conclusions and future work

We presented two novel algorithms for placement of VMs in data centers. Unlike existing research based on simple First Fit heuristics, our techniques take advantage of the polynomial-time linear and quadratic programs, and provide (near-)optimal solutions to the vector bin packing (VBP) problem.

Our experiments confirm the substantial improvement of our approach over the existing techniques and demonstrate that our algorithm ‘PACKINGVECTORS’ consistently yields the optimal placement across a broad spectrum of inputs. As part of future work, we intend to expand our techniques to handle dynamic placements and continuous optimization of data centers.

## 7 Acknowledgments

We would like to thank our colleagues for their help in the implementation of the ‘PACKINGVECTORS’ algorithm.

## References

- [1] Amazon EC2. <http://aws.amazon.com/ec2/>
- [2] GoGrid. <http://www.gogrid.com/>

- [3] Rackspace Cloud. <http://www.rackspacecloud.com/>
- [4] **lp-solve**. <http://lpsolve.sourceforge.net/5.5/>
- [5] N. Bobroff, A. Kochut, and K. Beaty. Dynamic Placement of Virtual Machines for Managing SLA Violations. *IM '07: Proceedings of the 10th IEEE Symposium on Integrated Management*, IEEE Computer Society, Washington, DC, USA (2007).
- [6] Nikhil Bansal, Alberto Caprara and Maxim Sviridenko. Improved approximation algorithms for multidimensional bin packing problems. *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society, Washington, DC, USA (2006) 697–708
- [7] M. Cardosa, M. R. Korupolu, and A. Singh. Shares and Utilities based Power Consolidation in Virtualized Server Environments. *IM '09: Proceedings of IFIP/IEEE Integrated Network Management*, (2009).
- [8] Nikhil Bansal, Jos R. Correa, Claire Kenyon and Maxim Sviridenko. Bin Packing in Multiple Dimensions: Inapproximability Results and Approximation Schemes. *Math. Oper. Res.* **31**(1), (2006) 31–49
- [9] R. Bianchini and R. Rajamony. Power and energy management for server systems. *IEEE Computer.* **37**, (2004) 2004.
- [10] Alberto Caprara. Packing 2-Dimensional Bins in Harmony. *FOCS '02: Proceedings of the 43rd Symposium on Foundations of Computer Science*, IEEE Computer Society, Washington, DC, USA (2002) 490–499
- [11] Alberto Caprara and Paolo Toth. Lower bounds and algorithms for the 2-dimensional vector packing problem. *Discrete Appl. Math.* **111**(3), (2001) 231–262
- [12] Chandra Chekuri and Sanjeev Khanna. On multi-dimensional packing problems. *SODA '99: Proceedings of the tenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, Baltimore, Maryland, USA (1999) 185–194
- [13] George Bernard Dantzig and Mukund Narain Thapa. *Linear Programming: Theory and extensions*, Springer-Verlag (2003).
- [14] Wenceslas Fernandez de la Vega and George S. Lueker. Bin packing can be solved within  $1+\epsilon$  in linear time. *Combinatorica* **1**(4), (1981) 349–355

- [15] M.R. Garey, R.L. Graham and D.S. Johnson. Resource Constrained Scheduling as Generalized Bin Packing. *J. Combinatorial Theory* **21**(3), (1976) 257–298
- [16] L. Grit, D. Irwin, A. Yumerefendi. and J. Chase. Virtual Machine Hosting for Networked Clusters: Building the Foundations for Autonomic Orchestration. *VTDC '06: Proceedings of IEEE International Workshop on Virtualization Technology in Distributed Computing*, IEEE Computer Society, Washington, DC, USA (2006).
- [17] Michael R. Garey and David S. Johnson. *Computers and Intractability: A guide to the Theory of NP-Completeness*, W.H. Freeman and Co. (1976).
- [18] F. Hermenier, X. Lorca, J.-M. Menaud, G. Muller and J. Lawall. Entropy: A Consolidation Manager for Clusters. *VEE '09: Proceedings of the 2009 ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments*, (2009).
- [19] Chris Hyser, Bret Mckee, Rob Gardner and Brian J. Watson. Autonomic Virtual Machine Placement in the Data Center. *HP Labs technical report*, HPL-2007-189. (2007)
- [20] David Karger and Krzysztof Onak. Polynomial approximation schemes for smoothed and random instances of multidimensional packing problems. *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2007) 1207–1216
- [21] Narendra Karmarkar and Richard M. Karp. An efficient approximation scheme for the one-dimensional bin-packing problem. *SFCS '82: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, IEEE Computer Society, Washington, DC, USA (1982) 312–320
- [22] M. K. Kozlov, S. P. Tarasov and L. G. Khachiyan. The polynomial solvability of convex quadratic programming. *USSR Computational Mathematics and Mathematical Physics*, **20**(5), (1980) 223–228
- [23] Hans Kellerer and Vladimir Kotov. An approximation algorithm with absolute worst-case performance ratio 2 for two-dimensional vector packing. *Operations Research Letters* **31**(1), (2003) 35–41

- [24] G. Khanna, K. Beaty, G. Kar and A. Kochut. Application Performance Management in Virtualized Server Environments. *NOMS '06: Proceedings of 10th IFIP/IEEE Network Operations and Management Symposium*, (2006) 373–381
- [25] K. G. Murty. Linear complementarity, linear and nonlinear programming. *Sigma Series in Applied Mathematics*, Berlin: Heldermann Verlag, (1988)
- [26] Boaz Patt-Shamir and Dror Rawitz. Vector Bin Packing with Multiple-Choice. *CoRR* **abs/0910.5599**, (2009) <http://arxiv.org/abs/0910.5599>
- [27] Alexander Schrijver. *Theory of Linear and Integer Programming*, Wiley-Interscience series in discrete mathematics, Wiley (1986).
- [28] Johara Shahabuddin, Abhay Chrungoo, Vishu Gupta, Sandeep Juneja, Sanjiv Kapoor and Arun Kumar. Stream-Packing: Resource Allocation in Web Server Farms with a QoS Guarantee. *HiPC '01: Proceedings of the High Performance Computing Conference*, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, (2001) 182–191
- [29] Vijay V. Vazirani. *Approximation Algorithms*, Springer-Verlag, Berlin (2001).
- [30] A. Verma, P. Ahuja, and A. Neogi. pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems. *IMC '08: Proceedings of the 9th ACM International Middleware Conference*, (2008).
- [31] W. Vogels. Beyond Server Consolidation. *ACM QUEUE* **6**(1), (2008) 20–26
- [32] Gerhard J. Woeginger. There is no asymptotic PTAS for two-dimensional vector packing. *Information Processing Letters* **64**(6), (1997) 293–297

## 8 Appendix

### 8.1 Dual formulation of the ILP

$$\text{minimize : } \sum_j y_j \quad (16)$$

$$\text{such that } \sum_j x_{ij} = 1 \quad 1 \leq i \leq n \quad (17)$$

$$\sum_i p_i^k \cdot x_{ij} \leq 1 \quad 1 \leq j \leq m, 1 \leq k \leq d \quad (18)$$

$$y_j \geq x_{ij} \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (19)$$

$$x_{ij} \in \{0, 1\} \quad 1 \leq i \leq n, 1 \leq j \leq m \quad (20)$$

Multiply constraint (19) by positive multipliers  $z_{ij}$  corresponding to  $x_{ij}$ 's. Adding all such constraints, we obtain -

$$\begin{aligned} (\sum_i z_{ij})y_j &\geq \sum_i x_{ij}z_{ij} & 1 \leq j \leq m \\ \sum_j (\sum_i z_{ij})y_j &\geq \sum_j \sum_i x_{ij}z_{ij} \end{aligned}$$

Further, we have -

$$\begin{aligned} \sum_j y_j &\geq \sum_j (\sum_i z_{ij})y_j \geq \sum_j \sum_i x_{ij}z_{ij} \\ \text{subject to } \sum_i z_{ij} &\leq 1 \end{aligned} \quad (21)$$

Thus, the minimization problem can be reframed as a maximization problem with the constraint (21) and objective function being -

$$\text{max : } \sum_i \sum_j x_{ij}z_{ij}$$

Adding the new constraints and relaxing constraint (20), the dual problem is as follows -

$$\begin{aligned}
& \text{maximize : } \sum_i \sum_j x_{ij} z_{ij} \\
\text{such that } & \sum_j x_{ij} = 1 & 1 \leq i \leq n \\
& \sum_i p_i^k . x_{ij} \leq 1 & 1 \leq j \leq m, 1 \leq k \leq d \\
& \sum_i z_{ij} \leq 1 & 1 \leq j \leq m \\
& x_{ij}, z_{ij} \geq 0 & 1 \leq i \leq n, 1 \leq j \leq m
\end{aligned}$$